A GENERALIZATION OF DIFFERENCE SETS[*]

Robert J. McEliece[+]

a. <u>Introduction</u>. A (v,k,λ) difference set D is a set of k distinct residues $\{a_1, a_2, \ldots, a_k\}$ modulo v such that every residue $b \not\equiv 0 \pmod v$ can be expressed in exactly λ ways in the form $b \equiv a_i - a_j \pmod v$. With each difference set we may associate a binary periodic sequence $(s_1, s_2, \ldots)$ with $s_i = 1$ if $i \pmod v$ is in D, and $s_i = 0$ otherwise. Since this sequence is periodic of period v, we need only consider one cycle from the sequence. Such cycles we agree to call (binary) difference cycles. Difference cycles (equivalently, difference sets) have been studied intensively (Refs. (1), (3)). They have important applications to digital communications, mainly because they have 2-level autocorrelation. In this paper we shall point out certain other (equivalent) properties of difference cycles which seem susceptable to immediate generalization, but show that these generalizations are vacuous.[**]

b. Motivation. If s is a difference cycle, then the defining property of difference sets tells us that the number of ordered pairs (s_i, s_{i+b}) from s (subscripts taken mod v) of the form (1,1) is λ for all values of $b \not\equiv 0 \pmod v$. More generally, let the number of ordered pairs (s_i, s_{i+b}) from s of the form (ϵ_1, ϵ_2) be denoted by $p_{\epsilon_1,\epsilon_2}(b)$. Thus $p_{1,1}(b) = \lambda$ for all $b \not\equiv 0 \pmod v$. A

[*] This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration.

[+] Communications Systems Research Section of the California Institute of Technology Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, California.

[**] The author wishes to thank Professor E. S. Selmer for suggesting this problem.

[redacted]

simple enumeration now shows that, in addition, $p_{0,1}(b) = p_{1,0}(b) = k - \lambda$, $p_{0,0}(b) = v - 2k + \lambda$, whenever $b \not\equiv 0 \pmod v$. For let us represent s and its b<u>th</u> translate s_b schematically as below:

$$\begin{array}{ll} s: & \overbrace{11\ldots1 \quad 11\ldots1}^{k} \quad \overbrace{00\ldots0 \quad 00\ldots0}^{v-k} \\ s_b: & \underbrace{11\ldots1}_{\lambda} \quad \underbrace{00\ldots0}_{k-\lambda} \quad \underbrace{11\ldots1}_{k-\lambda} \quad \underbrace{00\ldots0}_{v-2k+\lambda} \end{array}$$

Since exactly λ of the <u>ones</u> from s match up with 1's from s_b, the remaining $k-\lambda$ <u>ones</u> from s must be paired with <u>zeros</u> from s_b. This shows $p_{1,0}(b) = k-\lambda$. The other relations may be verified similarly.

Now let s be any binary cycle of length v (not necessarily associated with a difference set). The (unnormalized) autocorrelation of s, $R_s(b)$, has been defined as follows (Ref. (1)): $R_s(b) = A_s(b) - D_s(b)$, where $A_s(b)$ is the number of agreements between s and s_b (i.e., the number of components in which s and s_b have the same entry) and $D_s(b)$ is the number of disagreements. With the notation introduced above, $R_s(b) = p_{1,1}(b) + p_{0,0}(b) - p_{1,0}(b) - p_{0,1}(b)$. Our remarks then show in particular that for a difference cycle s, $R_s(b)$ is independent of b if $b \not\equiv 0 \pmod v$, and so s has two-level autocorrelation. (Of course $R_s(0) = v$). Conversely, it is easy to show that, except for <u>pulse cycles</u> (i.e., s = (1,0,0,...,0) or (0,1,1,...,1) or a translate of one of these), any cycle with two-level autocorrelation is associated with a difference set (Ref. (1)).

c. <u>Generalization</u>. The preceding discussion motivates the following formal

[redacted]

generalization of a difference cycle which depends neither on the notion of autocorrelation function nor on group theory. If $s = (s_1, s_2, \ldots, s_v)$ is any binary cycle, and if b is an integer satisfying $0 < b < v$, we define a *bigram* $M(s,b)$ as follows: $M(s,b) = \left\{(s_i, s_{i+b}): \ i = 1,2, \ldots, v,\right\}$, *multiplicity included*. We have seen that if s is a binary difference cycle, then $M(s,b) = M(s,b')$ (equality means that the two collections contain the same pairs with the same multiplicity) whenever $0 < b < v$, $0 < b' < v$. More generally, we define an m-gram for an n-ary cycle $s = (s_1, s_2, \ldots, s_v)$ as follows:

Definition: Let $\underline{b} = (b_1, b_2, \ldots, b_{m-1})$ be an ordered (m-1)-tuple of integers with $0 < b_1 < b_2 < \ldots < b_{m-1} < v$. We define the *m-gram* $M(s,\underline{b})$ to be the collection $\left\{(s_i, s_{i+b_1}, s_{i+b_2}, \ldots, s_{i+b_{m-1}}), \ i = 1,2, \ldots, v,\right\}$ multiplicity included. If $M(s,\underline{b}) = M(s,\underline{b}')$ for all such $\underline{b}$ and $\underline{b}'$, we say that s is *m-tuply regular*.

Thus in the new terminology an ordinary difference cycle becomes a doubly-regular binary cycle. We have formally generalized in two directions; we allow both the degree of regularity of s and the number of symbols in s to increase. We remark that m-tuply regular unary cycles and singly regular n-ary cycles are trivial, and that a pulse binary cycle is m-tuply regular in a trivial way. It is surprising that binary difference cycles and the above trivial examples are the only examples of m-tuply regulr n-ary cycles possible. We prove this result now.

d. *Non-Existence*.

Theorem 1: If s is an m-tuply regular n-ary cycle, then one of the four alternatives below holds:

(1) $m = 2$, $n = 2$ and s is a difference cycle

(2) $n = 2$ and s is a pulse

(3) $m = 1$

(4) $n = 1$.

The proof is in two parts: (1) We assume $m > 2$ and conclude $n = 1$ unless s is a pulse; (2) We assume $n > 2$ and show $m = 1$.

Suppose, then, that s is an m-tuply regular n-ary cycle with $m > 2$. It is clearly sufficient to prove this is impossible for $n = 2$, since if $n > 2$ we may identify certain of the symbols to obtain an m-tuply regular binary cycle. We now digress in order to place the problem in a wider context.

Definition: (Hananai, Ref. (2)). Let S be a set of v distinct objects. A tactical configuration $C = C\left[k,m,\lambda,v\right]$ is a collection of b subsets (called blocks) B_i, $i = 1,2, \ldots, b$, of S such that each block contains exactly k objects from S, and each (unordered) m-tuple from S occurs in exactly λ blocks. C is symmetric if $b = v$, and if each object in S occurs in exactly k blocks.

Our plan is to show that the existence of an m-tuply regular binary cycle implies the existence of a symmetric $C\left[k,m,\lambda,v\right]$. Theorem 2 shows that such configurations are trivial; we prove Theorem 2 first.

Theorem 2: There are no nontrivial symmetric $C\left[k,m,\lambda,v\right]$ configurations, if $m \geq 3$. (Trivial means that k does not satisfy $m \leq k \leq v - m$).

Proof: It is clear that a $C\left[k,m,\lambda,v\right]$ configuration is also a $C\left[k,m',\lambda,v\right]$ configuration for $m' \leq m$, since each unordered m'-tuple from S is a subset of exactly $\binom{v-m'}{m-m'}$ unordered m-tuples from S, and so each m'-tuple occurs

$\lambda\binom{v-m'}{m-m'} / \binom{k-m'}{m-m'}$ times in the configuration. Consequently it will be sufficient to prove theorem 2 for $m = 3$.

For the moment let $\lambda = \lambda_3$, and let λ_2 represent the number of times each unordered pair from S occurs in C. Then counting in two different ways the number of times a triple involving a given pair occurs in the configuration, we see that

$$\lambda_2(k-2) = \lambda_3(v-2) \quad . \tag{1}$$

Note that (1) holds for any $C\left[k,3,\lambda,v\right]$ configuration, symmetric or not.

If C is symmetric, let us count in two ways the number of times a pair involving a given element occurs:

$$k(k-1) = \lambda_2(v-1) \quad . \tag{2}$$

We now perform the standard trick (see Bose (5)) of deleting from C an arbitrary block, and all objects occuring in that block. Since C is in particular a symmetric block design, the derived design C' will also be a $C\left[k',3,\lambda',v'\right]$ configuration (but no longer symmetric) with $k' = k - \lambda_2$, $\lambda' = \lambda_3' = \lambda_3$, $\lambda_2' = \lambda_2$, $v' = v - k$. Eq. (1) will now apply to the derived parameters; i.e.,

$$\lambda_2(k - \lambda_2 - 2) = \lambda_3(v-k-2) \quad . \tag{3}$$

Combining Eq. (3) with Eq. (1), we see that $\lambda_2^2 / k = \lambda_3$. Thus $\lambda_2/k = (k-2)(v-2)$

and so, from Eq. (2), $(k-1)/(v-1) = (k-2)/(v-2)$; which implies $k = v$. But $k = v$ is a trivial design, and this completes the proof of Theorem 2.

To complete the first part of the proof of Theorem 1, it remains to show that the existence of an m-tuply regular binary cycle ($m \geq 3$) which is not a pulse implies the existence of a non-trivial symmetric $C\left[k,m,\lambda,v\right]$ configuration. First of all, it is clear that if $1 < k < m$, then no such cycle exists, since for certain b's, $M(s,b)$ will contain $(\overbrace{11 \ldots 1}^{k}\ \overbrace{00 \ldots 0}^{m-k})$, while others will not. Similarly $v - m > k > v - 1$ is impossible. Thus, except for pulses all m-tuply regular binary sequences with k 1's satisfy $m \leq k \leq v - m$.

If now s is an m-tuply regular binary cycle of length v (we assume the two symbols are 0 and 1), let $S = \left\{a_1, a_2, \ldots a_v\right\}$ be any set containing v distinct objects. We define blocks B_i, $i = 0,1,2, \ldots, v - 1$ as follows: $a_j \epsilon B_i$ if and only if $s_{i+j} = 1$. To show that these blocks form a symmetric $C\left[k,m,\lambda,v\right]$ configuration, we need only verify the m-tuple condition.

Thus, let $(a_{i_1}, a_{i_2}, \ldots, a_{i_m})$ be an m-tuple from S, and assume $i_1 < i_2 < \ldots < i_m$. Let $b_j = i_{j+1} - i_j$, $j = 1,2, \ldots, m - 1$, and set $\underline{b} = (b_1, b_2, \ldots, b_{m-1})$. Since s is m-tuply regular, the m-tuple $(1,1, \ldots, 1)$ will occur in $M(s,b)$ a certain number of times, say λ_m, and λ_m is independent of b. It is clear that if $(s_i, s_{i+b_1}, \ldots, s_{i+b_{m-1}})$ is such an m-tuple from $M(s,b)$, then B_{i-i_1} will contain the m-tuple $(a_{i_1}, a_{i_2}, \ldots, a_{i_m})$, and conversely. Hence every m-tuple from S occurs in exactly λ_m blocks, and so the blocks B_i

do form a (non-trivial) symmetric C $\left[k,m,\lambda,v\right]$ configuration. But this is impossible by Theorem 2, and so every m-tuply regular binary cycle is a pulse. This completes the first part of the proof of Theorem 1.

Our attention has recently been drawn to the fact that the second half of Theorem 1 was proved independently by R. Titsworth (Ref. 6) several years ago. (The reader who consults that report will see that Titsworth's "perfect" sequences are precisely the doubly regular sequences discussed here.) We present here a new proof which makes use of the highly-developed theory of difference sets.

In order to prove the second half of Theorem 1, we will assume that s is a doubly regular n-ary cycle, and show that $n > 2$ is impossible. It will be sufficient to prove that there are no doubly regular ternary cycles, since a doubly regular n-ary cycle $(n > 3)$ can be transformed into a doubly regular ternary cycle by a simple identification of certain symbols.

If s is a doubly regular ternary cycle in the symbols 0,1,2, let s contain k_0 zeros, k_1 ones and k_2 twos. We observe that each k_i must be ≥ 2, since if (say) $k_0 = 1$, then some bigrams would contain (0,1) but not (0,2), while others would contain (0,2) but not (0,1). Let us now identify the symbols 0 and 1; s then becomes a doubly regular binary cycle (which is not a pulse by the observation above), and so the set $D_2 = \left\{ i : s_i = 2 \right\}$ is a difference set. Similarly, D_0 and D_1 are also difference sets. But also $D_{0,1} = \left\{ i : s_i = 0 \text{ or } s_i = 1 \right\}$, being the complement of D_2, is a difference set. Also, $D_{0,2}$ and $D_{1,2}$, defined similarly, are also difference sets.

We remark at this stage that $D_0 \bigcup D_1 = D_{0,1}$, $D_0 \bigcap D_1 = \emptyset$. This means that if we could prove that the union of two disjoint difference sets is never a non-trivial difference set, the second part of Theorem 1 would follow as an immediate corollary. But although there is strong evidence that this is so (see the discussion at the end of this paper), we are as yet unable to prove it. So we must go another route.

To continue with our proof, write $D_0 = \{a_1, a_2, \ldots, a_{k_0}\}$ and $D_1 = \{b_1, b_2, \ldots, b_{k_1}\}$. We define the polynomials θ_0 and θ_1 (see Ryser Ref. (3)) as follows:

$$\theta_0(x) \equiv x^{a_1} + x^{a_2} + \ldots + x^{a_{k_0}} \pmod{x^v - 1}$$

$$\theta_1(x) \equiv x^{b_1} + x^{b_2} + \ldots + x^{b_{k_1}} \pmod{x^v - 1} \quad .$$

Then since D_0 and D_1 are difference sets, we have, as in Ref. (3),

$$\theta_0(x)\,\theta_0(x^{-1}) \equiv n_0 + \lambda_0\, T(x) \pmod{x^v - 1}$$

$$\theta_1(x)\,\theta_1(x^{-1}) \equiv n_1 + \lambda_1\, T(x) \pmod{x^v - 1} \quad , \tag{4}$$

where $n_0 = k_0 - \lambda_0$, $\lambda_0 = k_0(k_0 - 1)/(v - 1)$, $n_1 = k_1 - \lambda_1$, $\lambda_1 = k_1(k_1 - 1)/(v - 1)$, and $T(x) \equiv 1 + x + x^2 + \ldots + x^{v-1} \pmod{x^v - 1}$.

Now since s is doubly regular, the pair (1,0) occurs in each bigram M(s,b) equally often, say μ times. This says that every residue $b \not\equiv 0 \pmod v$ can be written in exactly μ ways in the form $b \equiv a_i - b_j \pmod v$. In terms of the Θ's, this condition becomes

$$\theta_0(x)\ \theta_1(x^{-1}) \equiv -\mu + \mu T(x) \pmod{x^v - 1} \quad . \tag{5}$$

If we multiply both sides of Eq. (5) by $\theta_1(x)$, and use Eq. (4), we obtain

$$n_1\theta_0(x) + \mu\theta_1(x) \equiv (\mu k_1 - \lambda_1 k_0)\ T(x) \pmod{x^v - 1} \quad . \tag{6}$$

(Observe that $R(x)T(x) \equiv R(1)T(x) \pmod{x^v - 1}$; see Ryser (3).) But the lefthand side of (6) cannot contain powers of x higher than $v - 1$, and so

$$n_1\theta_0(x) + \mu\theta_1(x) = (\mu k_1 - \lambda_1 k_0)\ T(x) \quad . \tag{7}$$

But this is impossible: the left-hand side of Eq. (7) cannot contain all powers of x less than v, since some s_i are equal to 2. This contradiction completes the proof of Theorem 1.

e. Conclusion. We remark finally that there is a certain amount of arbitrariness in our definition of multiple regularity. With hindsight at least, we might regard the fact that the pairs (0,1) and (1,0) occur evenly distributed among the bigrams of a difference cycle as a fluke, peculiar to the case of binary cycles. We could then define multiple regularity by only requiring that m-tuples of the form (aa ... a) be evenly distributed. If we had done this, the first part of the proof of Theorem 1 would still have worked, since we only needed the even distribution of (11 ... 1) anyway. But a new proof of the second part of the theorem would be needed; in fact, our conjecture that the union of two disjoint difference sets can never be a difference set is exactly what is required at least for ternary cycles. Some numerical evidence that this conjecture is true is available: for example, we require

$(v-1) \mid k_0(k_0-1)$, $(v-1) \mid k_1(k_1-1)$, $(v-1) \mid (k_0+k_1)(k_0+k_1-1)$, and Miller (Ref. 4) has found all triples (v, k_0, k_1) which satisfy these conditions, and with $k_0 \neq k_1$. It is easy to show that $k_0 = k_1$ is impossible. Using Miller's list, the author has been able to verify the conjecture for $v \leq 300$; there are 32 triples (v, k_0, k_1) in this range. There would appear to be no particular difficulty in pushing these numerical results even farther.

References

1. Golomb, et. al., Digital Communications With Space Applications, Prentice Hall, Englewood Cliffs, 1965.

2. H. Hananai, The Existence and Construction of Balanced Incomplete Block Designs, Ann. Math. Stat. 32, 1961, pp. 361-386.

3. H. J. Ryser, Combinatorial Mathematics, Carus Math. Monograph No. 14, Wiley, New York, 1963.

4. Miller: A partition problem, to appear.

5. R. C. Bose, On construction of balanced incomplete block designs, Ann. Eugen. 9 (1939), pp. 353-399.

6. R. Titsworth, Correlation Properties of Random-Like Periodic Sequences, J.P.L. Progress Report 20-391, October 1959.